

United States Patent Application

of

Johni Chan

HYBRID SWITCHING ARCHITECTURE

CERTIFICATE OF MAILING
BY "EXPRESS MAIL"

Express Mail Mailing Label No.

EK23A975316US


Date of Deposit:

22-MARCH-2001

I hereby certify that this paper is being deposited with the United States Postal Service "Express Mail Post Office to Addressee" Service under 37 CFR '1.10 on the date indicated above and is addressed to the Commissioner of Patents and Trademarks, Washington, D.C. 20231

THOMAS F. LEBENS

(Typed name of person mailing)



(Signature of person mailing)

HYBRID SWITCHING ARCHITECTURE

5

BACKGROUND OF THE INVENTION

The present invention relates to backup hardware in electronic computer systems, and, in particular, to backup hardware employing active single board computers (SBC's). Even more particularly, the present invention relates to a hybrid switching architecture for providing active-active single board computers backup in an electronic computer system.

During the past decade, the personal computer industry has literally exploded into the culture and business of many industrialized nations. Personal computers, while first designed for applications of limited scope involving individuals sitting at terminals, producing work products such as documents, databases, and spread sheets, have matured into highly sophisticated and complicated tools. What was once a business machine reserved for home and office applications, has now found numerous deployments in complicated industrial control systems, communications, data gathering, and other industrial and scientific venues. As the power of personal computers has increased by orders of magnitude every year since the introduction of the personal computer, personal computers have been found performing tasks once reserved to mini-computers, mainframes and even supercomputers.

In many of these applications, PC hardware and industry-standard software performs mission critical tasks involving significant stakes and low tolerance for failure. In these environments, even a single short-

lived failure of a PC component can represent a significant financial event for its owner.

Standard off-the-shelf computers and operating systems are used in critical applications that require much higher levels of reliability than provided by most personal computers. They are used for communications applications, such as controlling a company's voice mail or e-mail systems. They may be used to control critical machines, such as check sorting, or mail sorting for the U.S. Postal Service. They are used for complicated industrial control, automaton, data gathering and other industrial and scientific applications. Computer failures in these applications can result in significant loss of revenue or loss of critical information. For this reason, companies seek to purchase computer equipment, specifically looking for features that increase reliability, such as better cooling, redundant, hot-swappable power supplies or redundant disk arrays. These features have provided relief for some failures, but these systems are still vulnerable to failures of the single board computer (SBC) within the personal computer system itself. If the processor, memory or support circuitry on a single board computer fails, or software fails, the single board computer can be caused to hang up or behave in such a way that the entire computer system fails. Some industry standards heretofore dictated that the solution to this problem is to maintain two completely separate personal computer systems, including redundant single board computers and interface cards. In many cases, these interface cards are very expensive, perhaps as much as ten times the cost of the single board computer.

As a result, various mechanisms for creating redundancy within and between computers have been

attempted in an effort to provide backup hardware that can take over in the event of a failure.

In a typical computer system a common-bus architecture connects all components, which may include
5 one or several central processing units (CPUs), random access memory (RAM), read-only memory (ROM), input/output (I/O) devices, disk drive controllers, direct memory access controllers (DMAC), secondary bus controllers such as a small computer systems interface (SCSI) or bus
10 bridges to other buses such as a peripheral component interconnect (PCI), compact PCI, or an industry standard architecture (ISA) bus. Those components may all be disposed on a single plug-in board, or they may be implemented on a plug-in board as well as a motherboard.
15 In the later case, the plug-in board(s) and the motherboard communicate via a bus. In some cases, data is shared by multiple CPUs using multiple port memories or data must be accessed by various components, one component at a time, or transferred from one component to
20 another on a common bus.

The present invention advantageously addresses the above and other needs.

SUMMARY OF THE INVENTION

25 The present invention advantageously addresses the needs above as well as other needs by providing a hybrid switching architecture for providing connectivity among active switching single board computers in an electronic computer system.

30 In accordance with one embodiment, the present invention can be characterized as a system employing a first processor including a first processor data channel; a first hybrid switching module including a first hybrid switching module processor data channel, a first hybrid

switching module main data channel, a first input/output link data channel, and a first switch, the first hybrid switching module processor data channel being coupled to the first processor data channel; a first main bus
5 coupled to the first hybrid switching module main data channel; a second processor including a second processor data channel; and a second hybrid switching module including a second hybrid switching module processor data channel, a second input/output link data channel, and a
10 second switch, the second hybrid switching module processor data channel being coupled to the second processor data channel, the second input/output link data channel being coupled to the first input/output link data channel.

15 In accordance with another embodiment, the present invention can be characterized as a hybrid switching module having a hybrid switching module processor data channel; a hybrid switching module main data channel; an input/output link data channel; a switch
20 coupled to the hybrid switching module processor data channel; and a bridge coupled to the hybrid switching module main data channel; wherein the switch is selectively coupled to the bridge and selectively coupled to the input/output link data channel, wherein the hybrid
25 switching module processor data channel is thereby selectively coupled to the bridge and selectively coupled to the input/output link data channel.

BRIEF DESCRIPTION OF THE DRAWINGS

30 The above and other aspects, features and advantages of the present invention will be more apparent from the following more particular description thereof, presented in conjunction with the following drawings wherein:

FIG. 1 is a block diagram illustrating a hybrid switching architecture implemented in a computer system of one embodiment of the present invention;

FIG. 2 is a schematic block diagram illustrating the computer system such as in FIG. 1 including a modification of single board computer (system host) connections to PCI or CPCI backplane buses by rerouting signals from single board computers through respective hybrid switching modules;

FIG. 3 is a conceptual representation of bus connections between components of a computer system such as in FIG. 2;

FIG. 4 is an illustration of a hybrid switching module such as illustrated in the computer system of FIG. 2 comprising an input PCI or CPCI bus connection, an output PCI or CPCI bus connection, input/output links, and a crossbar switch and arbiter and a bridge; and

FIG. 5 is an illustration of a plurality of system hosts/hybrid switching module pairs, such as employ the hybrid switching module of FIG. 4, each coupled internally through a bridge to their respective PCI back plane busses, and each through an appropriate input/output link to each other hybrid switching module/system host pair.

Corresponding reference characters indicate corresponding components throughout the several views of the drawings.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

The following description of the presently contemplated best mode of practicing the invention is not to be taken in a limiting sense, but is made merely for the purpose of describing the general principles of the

invention. The scope of the invention should be determined with reference to the claims.

July 2
~~Referring first to FIG. 1, a block diagram is~~
shown illustrating a hybrid switching architecture
5 implemented in a computer system 100, an embodiment of
the present invention. Shown is a first single board
computer (SBC) 102 coupled through a hybrid switching
module (HSM) 104 to a PCI or CPCI backplane bus 106.
Coupled to the PCI or CPCI backplane bus 106 are a
10 plurality of PCI or CPCI peripheral slots 108, into which
PCI or CPCI peripheral cards can be inserted, and user
CPCI year input/output devices 110. Also coupled to the
hybrid switching module 104 is an I/O link 112, which
couples to a second hybrid switching module 114. The
15 second hybrid switching module is connected to a second
single board computer (SBC) 116, and also to a second PCI
or CPCI backplane bus 118. Coupled to the second PCI or
CPCI backplane bus 118 are a second plurality of PCI or
CPCI peripheral slots 120, into which PCI peripheral
20 cards (not shown) can be inserted, and second user CPCI
~~rear input/output devices 122.~~

In practice, the hybrid switching module 104
(the first hybrid switching module 104) and the second
hybrid switching module 114 manage a flow of data between
25 the single board computer 102 (first single board
computer 102) and the second single board computer 116,
respectively. The single board computer (first single
board computer 102) and the second single board computer
116 have no control, at least directly, over their
30 respective abilities to access their respective PCI or
CPCI backplane buses 106, 118. Through the I/O link 112,
the first single board computer 102 and the second single
board computer 116 can be coupled through the first
hybrid switching module 104 and the second hybrid

switching module 114, to any other hybrid switching module within the computer system 100 and in turn coupled to any other PCI or CPCI backplane bus in the computer system 100.

5 Arbitration/switching is handled in the crossbar switch inside the hybrid switching module (HSM) and is controlled by middleware (e.g., cluster software, such as is widely available) executed by the single board computer that directs the hybrid switching module via the
10 crossbar switch. The bridge 410 (FIG. 4) translates signals from the single board computer into signals on the PCI or CPCI bus.

As a result, all of the PCI peripheral slots 108, 120, and the user rear PCI input/output devices 110, 122 are selectively available to each of the single board
15 computers 102, 116. As a result, the single board computers 102, 116 may both operate together in order to perform multiprocessor functions, or one of the single board computers 102, 116 may serve as a backup to the
20 other of the single board computers 102, 116, while not requiring full redundancy of the PCI peripheral cards in the PCI peripheral slots 108, 120 or the user PCI input/output devices 110, 122. This multi-computing backup can be extended to multiple single board
25 computers, e.g., up to 8 single board computers, inter or intra box connectivity as shown in FIG. 5.

~~Because each of the single board computers 102, 116 is able to connect to each of the PCI Peripheral slots 108, 120 and each of the user PCI input/output
30 devices 110, 122 in the tremendous amount of functionality is available to the user of the computer system. For example, in an eight-way multi-computing configuration, eight-way point-to-point connectivity and redundancy is enabled by the present embodiment. The~~

same connectivity can be applied directly as I/O chassis connectivity as the industry migrates into point-to-point architecture.

Sub 4
 5 ~~As shown by way of illustration, two single~~
 board computers, two hybrid switching modules, two PCI backplane busses, two sets of PCI peripheral slots and two sets of user PCI input/output devices are employed in the present embodiment. However, as will be appreciated by a person of ordinary skill in the art, the teachings
 10 at the present embodiment can be expanded to any member N of such components, wherein the I/O link 120 links all of the hybrid switching modules together, and allows each of N single board computers to communicate with each of N PCI backplane buses and each of the other single board
 15 computers via respective pairs of hybrid switching modules.

In this way an unlimited degree of redundancy (N times redundancy) can be achieved using the hybrid switching module of the present embodiment.

20 Referring next to FIG. 2, shown is a schematic block diagram illustrating the computer system including a modification of single board computer (system host) connections to PCI or CPCI backplane buses 206, 208 by rerouting signals from single board computers through
 25 respective hybrid switching modules 214, 216.

Sub 5
 30 ~~Shown is a single board computer coupled to a~~
 first hybrid switching module 214, and a second single board computer 212 coupled to a second hybrid switching module 216. The first hybrid switching module 214 and the second hybrid switching module 216 are coupled together through an I/O link 218 and are each coupled to respective PCI backplane buses 206, 208. The respective PCI backplane buses 206, 208 are coupled respectively to

~~PCI peripheral slots and user PCI or CPCI input/output devices 220, 222.~~

As shown in FIG. 2, a connection 224 between the first PCI backplane bus 206 and the first single board computer 210 is deleted, i.e., the central resources for the single board computer (the first single board computer 210) are disabled, so as to require that any access, i.e., communication/data transfer, between the first PCI backplane bus 206 and the first single board computer 210 be redirected through the first hybrid switching module 214. Similarly, a connection 226 between the second PCI backplane bus and the second single board computer 212 is deleted, i.e., eliminated, so as to require that any access, i.e., communication/data transfer, between the second PCI backplane bus 208 and the second single board computer 212 be redirected through the second hybrid switching module 218. Furthermore, the first hybrid switching module 214 and the second hybrid switching module 216 are coupled to each other through the I/O link 218.

This can be expanded to N connections through the I/O link 218 with arbitration being handled in the crossbar switch in each hybrid switching module, managed and controlled by the middleware.

Referring next to FIG. 3, illustrated is a conceptual representation of bus connections between components of a computer system 300 in accordance with the present embodiment.

As illustrated, a first single board computer 302 and a second single board computer 304 are coupled through bus connectors P3 and P5 314, 316 to a first hybrid switching module 306 and a second hybrid switching module 308, respectively. The hybrid switching modules in turn carry data from the bus connections P3 and P5

314, 316 to bus connections P1 and P2 318, 320, the standard bus connections from PCI backplane buses to single board computers. The hybrid switching modules 306, 308 carry the data from the single board computers 5 302, 304 to respective PCI backplane buses, which carry the data to respective PCI peripheral slots and user PCI input/output devices 310, 312.

Referring next to FIG. 4, illustrated is a hybrid switching module 400 comprising a PCI CPU bus connection 402, a PCI main bus connection 404, 10 input/output links 406, and a crossbar switch and arbiter 408 (such as for example INTEL Part No. 82808) and a bridge 410 (such as for example INTEL Part. No. 21554). (Alternatively, the hybrid switching module may be 15 implemented as a single integrated device, such as a single chip integrated circuit, or a multi-chip module.)

Data received through the PCI backplane bus 202 (FIG. 2), 204 (FIG. 2) is received in the PCI main bus connection 404 and directed to the bridge 410. The 20 bridge 410 directs the data from the PCI main bus connection 404 to the crossbar switch and arbiter 408, and performs other bridge functions, such as are well understood by the person of ordinary skill in the art.

The PCI CPU bus connection 402 is coupled 25 directly to the respective single board computer 210 (FIG 2), 212 (FIG. 2).

The PCI main bus connection 404 is coupled to the respective PCI backplane bus 202 (FIG. 2), 204 (FIG. 2).

30 In operation, the crossbar switch and arbiter 408 controls data flow between the PCI CPU bus connection 402 and the input/output link 406, and between the PCI CPU bus connection 402 and the PCI main bus connection 404. The data flow is controlled by software depending

on the mode of operation. In the case of CPU failure, for example, a failover operation will take place, the next available CPU will take over and be coupled to the PCI backplane busses through the hybrid switching modules. Thus, a means of data flow is provided between both the single board computer 210 (FIG. 2), 212 (FIG. 2), and each respective PCI backplane bus 202 (FIG. 2), 204 (FIG. 2) and each other single board computer 210 (FIG. 2), 212 (FIG. 2) and each PCI backplane bus 202 (FIG. 2) (via the I/O link 406).

Referring next to FIG. 5, illustrated is a plurality of single board computer/hybrid switching module pairs 500, 502, 504, 506, each coupled internally through a bridge to their respective PCI backplane buses 508, 510, 512, 514 and each through an appropriate input/output link 516, 518, 520, 522, 524, 526 to each other single board computer/hybrid switching module pair 500, 502, 504, 506. As shown, all of the PCI backplane buses 508, 510, 512, 514 may be part of an I/O chassis 528 that houses all of the point-to-point linkages (PCI backplane busses) 508, 510, 512, and 514. Each single board computer/hybrid switching module pair 500, 502, 504, 506 can be coupled to each of the illustrated point-to-point linkages 508, 510, 512, 514, either directly, through its own hybrid switching module and internal bridge or indirectly through one of the input/output links 516, 518, 520, 522, 524, 526 to one of the other hybrid switching modules and through the other hybrid switching module's internal bridge forming a high speed switching fabric.

Functioning of the hybrid switching modules 400 (FIG. 4) and in particular the operation of the crossbar switch and arbiter 408 (FIG. 4), in order to direct data from the respective single board computers to the

appropriate PCI backplane bus or other single board computer may be programmatically controlled, by, for example, the single board computers or may operate as a result, for example, of heartbeats that detect failure of various components within the industrial computer system switch out single board computers in the event failure in the single board computer is detected.

Advantageously, in the event of the failure of a single board computer, reduction in performance need only be reduced by, for example, 25 percent (in the event there are four single board computer/hybrid switching module pairs) as the hybrid switching modules can be reconfigured in order to utilize, for example, the remaining three single board computers.

While the invention herein disclosed has been described by means of specific embodiments and applications thereof, numerous modifications and variations could be made thereto by those skilled in the art without departing from the scope of the invention set forth in the claims.

By way of example, The HSM can be implemented as a rear I/O module and bridges to active central resources or it can be implemented as a companion board attached to the single board computer and bridges to active central resources.